

A Bayesian optimisation procedure for estimating optimal trajectories in electromagnetic compliance testing

Rémi Delanghe ^{1,†}, Tom Van Steenkiste ^{2,†}, Ivo Couckuyt ^{3,†,*}, Dirk Deschrijver ^{4,†} and Tom Dhaene ^{5,†}

Ghent University, 9000 Gent, Belgium; Remi.Delanghe@UGent.be (R.D.); tom.d.vansteenkiste@ugent.be (T.V.S.); dirk.deschrijver@ugent.be (D.D.); Tom.Dhaene@UGent.be (T.D.)

* Correspondence: Ivo.Couckuyt@UGent.be

† Current address: Technologiepark-Zwijnaarde 126 Ghent 9052, BELGIUM

Received: 30 August 2020; Accepted: 21 October 2020; Published: 30 October 2020

Abstract: The need for accurate physical measurements is omnipresent in both scientific and engineering applications. Such measurements can be used to explore and characterize the behavior of a system over the parameters of interest. These procedures are often very costly and time-consuming, requiring many measurements or samples. Therefore, a suitable data collection strategy can be used to reduce the cost of acquiring the required samples. One important consideration which often surfaces in physical experiments, like near-field measurements for electromagnetic compliance testing, is the total path length between consecutively visited samples by the measurement probe, as the time needed to travel along this path is often a limiting factor. A line-based sampling strategy optimizes the sample locations in order to reduce the overall path length while achieving the intended goal. Previous research on line-based sampling techniques solely focused on exploring the measurement space. None of these techniques considered the actual measurements themselves despite these values hold the potential to identify interesting regions in the parameter space, such as an optimum, quickly during the sampling process. In this paper, we extend Bayesian optimization, a point-based optimization technique into a line-based setting. The proposed algorithm is assessed using an artificial example and an electromagnetic compatibility use-case. The results show that our line-based technique is able to find the optimum using a significantly shorter total path length compared to the point-based approach.

Keywords: line-based sampling; Bayesian optimisation; exploitative sampling

1. Introduction

In many fields of science and engineering, measurements are an essential way to acquire a representation of an underlying, often complex, function. Determining a sampling strategy is often done through Design of Experiment (DoE) methods [1,2]. These methods aim at optimizing the locations of the samples such that the information gain per sample is maximized. There is a clear trade-off between the amount of samples taken and the associated cost of the experiment. Increasing the amount of samples can increase the quality of the representation, but will also increase the cost of the complete set of experiments.

When measurements are taken in a virtual space, such as a parameter space in a simulation, there is no constraint on the locations of sequentially taken samples. In a physical measurement space, using a measurement probe, the probe has to physically move and hence all samples form a path

with a predetermined order. This is referred to as a Line-based sampling strategy. Examples include Electromagnetic compatibility (EMC) testing [3], organ palpation [4] and vegetation monitoring [5]. The length of this path has to be taken into account in order to minimize sampling cost, as moving the probe takes measurement time, possibly resulting in an economic cost. The path length might also be limited due to physical constraints, for instance battery lifetime or the chance of mechanical failure [6]. However, as sequentially taken samples are close together, the additional information of a subsequent sample is not optimal. Here, a new trade-off emerges. On the one hand, the path between samples should be as short as possible to minimize the amount total travel time, and on the other hand, the space between samples should be as large as possible, which increases the path length.

We build further on Bayesian optimization (BO) to account for the cost of traveling between consecutively visited samples by creating a line-based extension to the framework. We solve an EMC problem and show that our technique has a better optimisation performance with respect to the path length.

This paper is structured as follows. First, in section 2 some background information is given regarding the literature of design of experiments. Next, in section 3 Bayesian optimisation is explained and in subsection 3.4 line-based sampling using BO is elaborated upon. Next a set of experiments comparing our approach to traditional solutions are performed in section 4, after which conclusions are drawn in section 5.

2. Background

The goal of any Design Of Experiment (DoE) method is to collect a data-set which covers the measurement space in an efficient way according to the task and goal at hand. For instance, in an exploration context, an adequate DoE could uniformly sample the measurement space. However, in a maximisation context, the DoE should focus on areas in the measurement space which have a high probability of containing the the maximum of the space. In the remainder of this section, we distinguish DoE methods in a number of properties.

2.1. Line-based Versus Sample-Based

The distinction is made between strategies which only focus on the sample locations, *point based* strategies, and the strategies which also focus on the path between the samples called *line based* strategies. A DoE method is classified as point based when it does not take into account the cost of movement between sample points. This is, for example, the case for sampling methods tailored to a virtual measurement space such as in computer simulations. Some examples include factorial designs [7] and FLOLA-voronoi sampling [8,9].

In line-based sampling, the cost of moving around in the measurement space is taken into account. All sample points are taken in a specific order and form a path. The information gain per sample should be maximized while keeping the path length to a minimum. Line-based sampling is very similar to coverage path planning [10], where the measurement probe must pass through every point in the measurement space. Examples include the Boustrophedon path [11] and the Hilbert curve [12].

2.2. Sequential Versus One-Shot

When a fixed number 'N' of sample points is allocated to the DoE method, it is characterised as a one-shot method. An example is factorial designs [7]. Determining the amount of samples upfront, allows for an ideal placement and coverage of the space. However, determining the optimal amount of samples N can be difficult and underestimations or overestimations can lead to additional costs. These types of technique are not adaptive as they do not allow for the DoE to be extended once it has been executed.

Sequential methods on the other hand, iteratively increase the amount of samples by determining consecutive sample locations based on the location and information gained from the previous samples. Theoretically, these algorithms could continue indefinitely. In practice, stopping criteria are used to determine when sufficient samples are available. Examples of sequential DoE include the point based maximin sampling [13] and the line-based Albatros algorithm [14]. The iterative process makes sequential methods more desirable in situations where it is uncertain how many samples will be needed before a certain criteria is reached or where mechanical failure of the measurement probe is likely.

2.3. Exploitative Versus Explorative

When the main goal of a DoE is to uniformly distribute samples within the measurement space, an explorative strategy is most suitable. In these strategies, only the measurement location and not the measured sample value is taken into account. Examples include the Hilbert curve [12] and the Albatros algorithm [14].

This is in contrast to exploitative DoE methods that use the measured sample values in order to optimize the location of upcoming samples and to focus on a region of interest, like for example the maximum of the sampled function. It is evident that exploitative methods are sequential by nature. This is because the information of previous sample values have to be taken into account before new samples can be taken. An example of a point-based sequential technique is the efficient global optimization (EGO) algorithm [15] using bayesian optimisation (BO).

It is important to note that the distinction between explorative and exploitative methods is not binary. An exploitative algorithm should encompass a degree of exploration in order to increase the knowledge about the environment which then can then be exploited. This trade-off between exploration and exploitation is a more general problem in the field of surrogate modeling and artificial intelligence [16–18] and in other fields like organisational management [19].

To the authors knowledge, no generic line-based sequential exploitative DoE method exists. This paper aims at filling in this gap as demonstrated in Table 1. The advantage of such a technique is that it is able to explore the sampling space but also that it is able to exploit the taken measurements, resulting in a swifter achievement of the set goal.

Table 1. Classification of sample strategy examples. The distinction is made between line-based and point-based, sequential and one-shot, and exploitative and explorative. The Line-Based BO sampling strategy fills in the gap of sequential exploitative line-based sampling strategies.

Examples	Line-based vs Point-based	Sequential vs One-shot	Exploitative vs Explorative
Factorial Designs [7]	Point-based	One-shot	Explorative
Boustrophedon path [11] Hilbert curve [12]	Line-based	One-shot	Explorative
Maximin criterion [13]	Point-based	Sequential	Explorative
ALBATROS [14].	Line-based	Sequential	Explorative
EGO [15] FLOLA-voronoi [8,9]	Point-based	Sequential	Exploitative
LBO (This work)	Line-based	Sequential	Exploitative

3. Bayesian Optimisation

3.1. Gaussian Process

In order to mathematically reason over the measurement space of the Device Under Test (DUT) given a finite set of samples, a surrogate model is built. This model mimics the response over the measurement space and can be used to perform cheap evaluations. The flexibility of the Bayesian optimisation framework allows for the use of a wide range of mathematical models. In this paper, the Gaussian Process (GP) is chosen as it has a high modeling power in a sparse-data context. A GP models the measured quantity over the DUT as a distribution over functions $f : X \rightarrow \mathbb{R}$. Given a finite set of input samples it is able to assign an estimate and uncertainty measure of the measured quantity at any location of the DUT. A GP is completely defined by a mean function $m : X \rightarrow \mathbb{R}$ and a covariance function $k : X \times X \rightarrow \mathbb{R}$ and the set of samples. In practice, the mean function is often considered to be 0, meaning that the only necessary design choice is the kernel function. For engineering applications, a common choice is the Matérn covariance function or the RBF covariance function. The latter is used in this paper and is given by equation 1. Here, σ is a parameter which is optimised during the learning process.

$$K(\mathbf{x}, \mathbf{x}') = \exp \left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\sigma^2} \right) \quad (1)$$

For a more in depth discussion about the use of GP's in machine learning, [20] can be consulted.

3.2. Optimisation Procedure

Given an adequate model, the BO procedure consists of multiple steps.

1. First the measured quantity is evaluated on an initial set of locations of the DUT.
2. This allows for building the GP model, over which an objective function $\alpha : X \rightarrow \mathbb{R}$ is defined and referred to as the acquisition function which is discussed in more detail in the next section.
3. As the GP is much cheaper to evaluate, it can be used to efficiently optimise $\alpha(x)$. The optimum x^* of $\alpha(x)$ is then evaluated on the actual DUT.
4. This evaluation produces a new sample which can be fed into the GP model.

By iterating over the above steps until a stopping criteria is reached, the model is iteratively refined in areas where the real optimum on the DUT most likely lies. This increases the chance of finding the real optimum of the true problem. This process is illustrated in Figure 1 by following the red path. This is also illustrated in Algorithm 1.

Algorithm 1: Bayesian Optimisation Process

```
// Initial dataset that is evaluated on the Device Under Test
Dataset = (x, y)
while End Condition Not Met do
    Build Model Using Dataset
    Construct the acquisition function  $\alpha : X \rightarrow \mathbb{R}$ 
     $x^* = \operatorname{argmax}_x \alpha(x)$ 
    Add  $(x^*, DUT(x^*))$  to Dataset
end
```

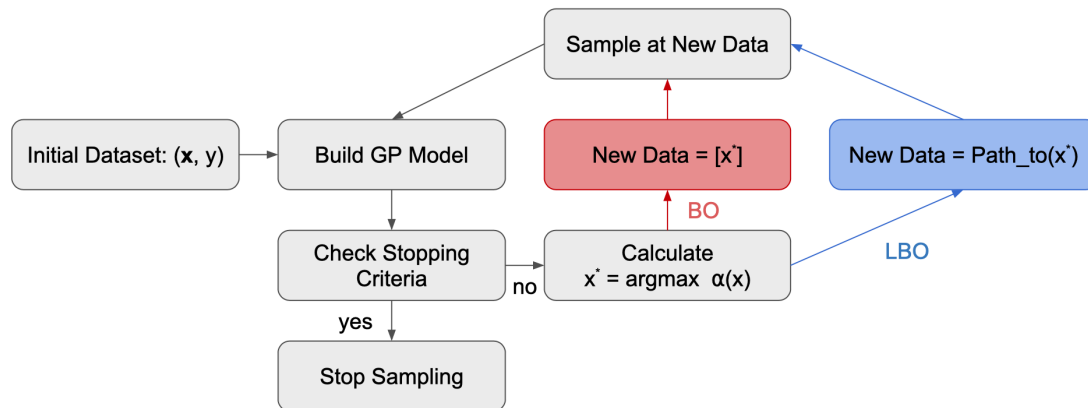


Figure 1. Illustration of the BO process and the adapted step for the line-based approach.

3.3. Acquisition Function

In order to find the next sampling location, an objective function, also called the acquisition function, is optimised over the GP model. For optimisation of the measured quantity, typical acquisition functions are Expected improvement (EI), Probability of Improvement (POI) [2] and Max-Value Entropy Search (MES) [21]. For more complex objectives a custom acquisition function can be constructed. In this paper the EI acquisition function is used.

$$EI(x) = (\mu(x) - f^*)\Phi\left(\frac{\mu(x) - f^*}{\sigma(x)}\right) + \sigma(x)\phi\left(\frac{\mu(x) - f^*}{\sigma(x)}\right) \quad (2)$$

As can be seen from the formula, the function uses both the mean estimate $\mu(x)$ and the variance $\sigma(x)^2$ from the Gaussian process to assess the quality of a sample at a location x . In addition, this acquisition function uses f^* , Φ and ϕ , which respectively represent the already maximum sampled value, the cdf of the normal distribution and the pdf of the normal distribution.

The acquisition function represents the improvement that is expected over the current best result if the next sampling point is at location x . The function offers a natural trade-off between exploration and exploitation as can be seen in its equation: the first term biases the search towards regions of interest and the second term towards regions of uncertainty [22].

3.4. Line-based Bayesian Optimisation

We propose to extend the Bayesian optimisation procedure by creating the Line-Based Bayesian Optimisation (LBO). This is illustrated by the blue path in Figure 1.

The procedure follows the BO strategy closely. It also starts by building a model using an initial dataset, which can be used to define the acquisition function $\alpha(x)$. The acquisition function is also optimised as discussed in the BO case, however instead of evaluating the model directly at x^* , a path of multiple samples is constructed from the current location of the probe to x^* . This results in more samples for a similar total path length. All the samples are fed to the GP model. This results both in a better

exploration of the measurement space as well as an increase in its predictive power. This procedure is also described in Algorithm 2

Algorithm 2: Line-based Bayesian Optimisation Process

```
// Initial dataset that is evaluated on the Device Under Test
Dataset = (x, y)
while End Condition Not Met do
    Build Model Using Dataset
    Construct the acquisition function  $\alpha : X \rightarrow \mathbb{R}$ 
     $x^* = \operatorname{argmax}_x \alpha(x)$ 
    Path = Path(from: current_location, to:  $x^*$ )
    Add Path to Dataset
end
```

The shape of the path from the current location to x^* will have an influence on the performance of the algorithm. In this work, a straight line is used with a set step-size. As this path is fully defined by a starting point and end point or better, the current location and the optimum of the acquisition function x^* , no extra optimisation step is needed. The step-size is a parameter to be set by the domain expert and should be set such that the function between two consecutively chosen points is expected to be well-behaved. Future work might explore more generic path shapes and adaptive step-size to further increase the performance of the algorithm.

4. Experiments

The LBO strategy is compared to the classic BO approach, using both the Expected improvement acquisition function and the RBF covariance function. The LBO uses a step-size of 0.1 units. Both techniques are initialised with one sample in the lower left corner of the sampling space. This mimics the EMC measurement probe entering the measurement space. All techniques are run for 30 iterations. In order to deal with varying final path lengths in different runs, the metrics are extrapolated using constant extrapolation. Due to the stochastic nature of the optimisation process used in the training of the GP model and the acquisition function, slight variations between runs are expected. For this reason, each experiment is repeated 15 times.

4.1. Use Cases

The techniques are tested on two 2D benchmark functions and a 2D engineering use case. The peaks benchmark function can be found as built-in function in Matlab® and is also used in [8], see Figure 2a. The formula is given by,

$$\begin{aligned} f(x, y) = & 3(-1 - 4x)^2 \times \exp(-((4x - 2)^2) - (4y - 1)^2) \\ & - 10((1/5) \times (4x - 2) - (4x - 2)^3 - (4y - 2)^5) \times \exp(-(4x - 2)^2 \\ & - (4y - 2)^2) - (1/3) \times \exp(-(4x - 1)^2 - (4y - 2)^2) \end{aligned} \quad (3)$$

The second benchmark function is a scaled and translated version of the peaks such that it presents a more challenging use case due to the large uniform area it creates. This function is represented in Figure 2b. The 2D engineering use case is an EMC measurement use case visualised in Figure 2c.

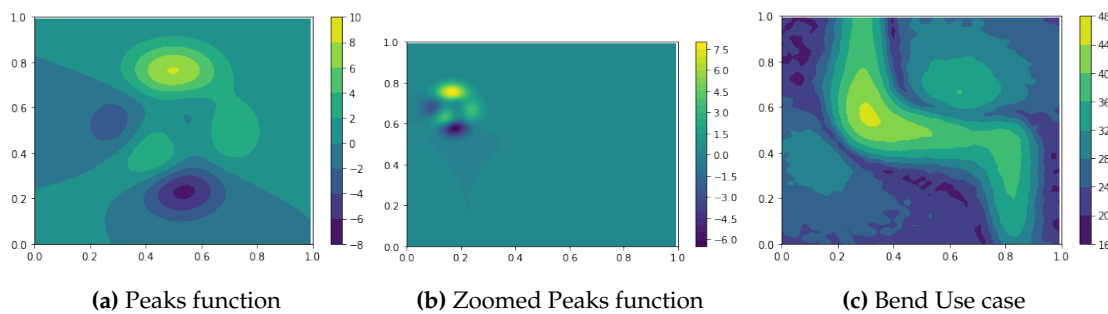


Figure 2. Visualisation of the use cases.

4.2. Evaluation Metrics

The optimisation performance of the algorithms is measured using regret. In addition, the Root mean squared error (RMSE) and the Mean Variance (MV) give additional insight into the quality of the fit. The first metric by which the technique is tested is Regret as described in Equation 4.

$$\text{REGRET} = \max_i (f(\text{sample}_i)) - \max_{x,y} (f(x,y)) \quad (4)$$

This metric measures the ability of the strategy to find the optimum in the measurement space. It compares the maximum of all samples taken by iteration i with the actual maximum over the complete measurement space.

The second metric is the RMSE between the prediction of the model and the ground truth. This measures the ability of the underlying GP to model the measured quantity in the measurement space by evaluating the mean estimation of the GP (μ) and the actual function over a finite set of i sample locations.

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\mu(x_i, y_i) - f(x_i, y_i))^2} \quad (5)$$

Finally, the MV metric measures the estimated uncertainty of the GP with respect to its modeling of the measured quantity. This measure can be calculated because the GP model inherently has uncertainty quantification (σ) from which one can sample over the measurement space.

$$\text{MV} = \frac{1}{N} \sum_{i=1}^N \sigma(x_i, y_i) \quad (6)$$

The three metrics are measured with respect to the pathlength of the measurement strategy. For each of the strategies, a lower value of the Regret indicates a better optimisation performance, a lower value of the RMSE indicates a better fit of the model to the data and a lower value of the MV indicates a higher confidence of the model in its predictions.

4.3. Results

Figure 3 illustrates resulting paths of both approaches and show the advantage of the LBO approach. Although less samples are taken in the classic BO approach, than using the LBO approach, the sample space is better explored, using a shorter total path length.

The BO strategy starts by exploring the space and by sampling at the borders. However, the samples are distant from each other and as such, a long path is needed to connect only a few samples. The strategy

does not take advantage of the fact that the probe passed over some regions and misses on the opportunity to capture information at that time.

The LBO strategy however, does sample whilst traveling between consecutive x^* locations. This means that the amount of data captured with respect to the total path length is much larger. This allows for better modelling of the measured quantity resulting in a quicker and more robust detection of the maximum value in the measurement space.

This is further backed by the full results of the experiments. Figures 4–6 show the evolution of the Regret, RMSE and MV metric for the different sampling strategies in the different use cases. In all use cases, the LBO strategy is able to discover the optimum of the sampling space significantly quicker than the classic BO version. For the peaks use case, LBO discovers the optimum twice as fast as BO (3.5 length units versus 7.5 length units).

The second use case, the zoomed peaks example, also shows that the LBO strategy outperforms the BO strategy with respect to the regret metric, however in this use case both strategies are not robust and show a high variability in the results over multiple runs. The lack of improvement of the RMSE and the low MV of both strategies before 4 length units shows that the model is overconfident and shows that the GP experiences difficulties modelling the highly flat regions in the space.

In the EMC use case, the LBO strategy is able to find the optimum using a path of 4 length units whilst the BO strategy, on average is not able to find it in double the path length. In addition to finding the optimum using a shorter path, the results are more robust against the inherent randomness of the optimisation process as the standard deviation of the LBO strategy are much smaller than the BO strategy. For these two use cases, the model of the measured quantity is also more accurate in the LBO strategy than the BO strategy. The increased amount of samples allows for a better model of the measured quantity and as such allows for increased predictive power, ie., a lower RMSE and for increased confidence ie., a lower MV, which can lead to a better estimation of the maximum location in the measurement space.

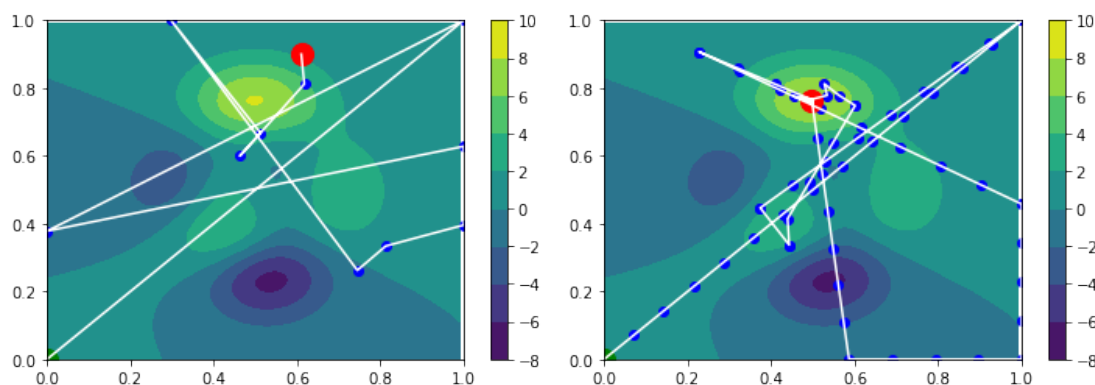


Figure 3. Example of a path taken by the BO and LBO strategy with similar total path length.

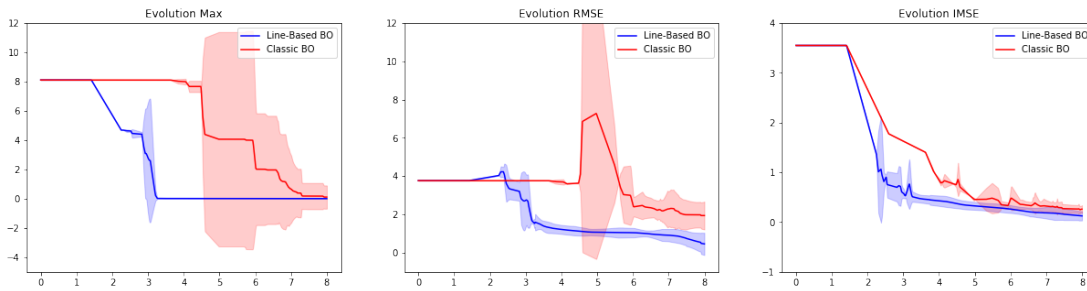


Figure 4. Comparison of the Mean and $2 \times \text{std}$ margins for the evolution of the Regret, RMSE and MV metric with respect to the path length for the peaks use case between the BO and LBO approach.

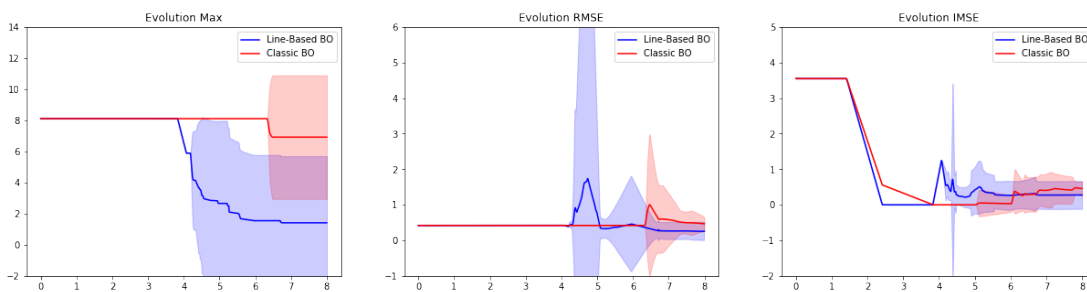


Figure 5. Comparison of the Mean and $2 \times \text{std}$ margins for the evolution of the Regret, RMSE and MV metric with respect to the path length for the zoomed peaks use case between the BO and LBO approach.

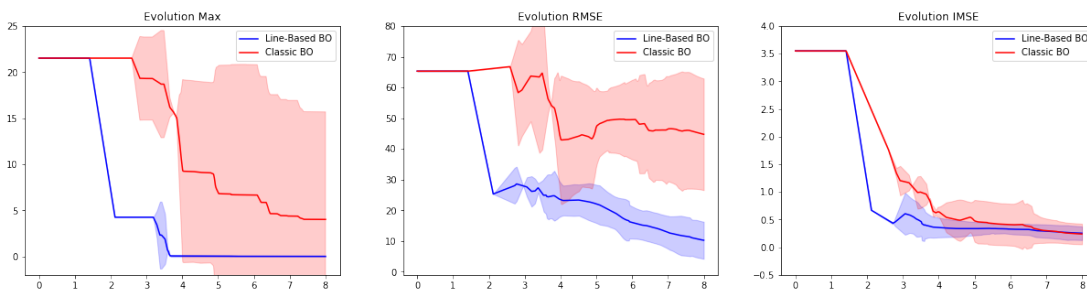


Figure 6. Comparison of the Mean and $2 \times \text{std}$ margins for the evolution of the Regret, RMSE and MV metric with respect to the path length for the EMC use case between the BO and LBO approach.

5. Conclusion

In this paper, the LBO Sampling strategy was introduced, benchmarked and applied to an existing engineering problem. The aim of the algorithm is to fill in the gap in DoE algorithms with a line-based, sequential and exploitative algorithm. The algorithm uses a Gaussian Process as underlying model to define waypoints between which to travel and to construct a path between the waypoints. The use of the Gaussian Process and a suitable acquisition function allows for balancing exploration and exploitation within the algorithm.

Whereas point based strategies optimize a metric in terms of the number of samples, line-based strategies focus on spreading the amount of samples while optimizing the path length. The LBO sampling strategy augments the capabilities of line based sample strategies by dynamically focusing on regions

of interest and on regions with limited knowledge. For both the benchmark experiment and the EMC engineering use-case the use of LBO shows a significant improvement to capture the underlying function and discover the global maximum while decreasing the total path length. This both increases the speed and decreases the cost of the sampling process which makes the LBO sampling strategy especially interesting for time consuming experiments as well as in situations where the amount of samples is not known upfront or where failure of the measurement probe is possible.

Future work might explore more generic path shapes and adaptive step-size to further increase the performance of the algorithm.

Funding: This work is partially supported by the Flemish Government under the “Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen” programme

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kleijnen, J.P. *Design and analysis of simulation experiments*. International Workshop on Simulation. Springer: Manhattan, New York City, 2015, pp. 3–22.
2. Snoek, J.; Larochelle, H.; Adams, R.P. Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems*, 2012, pp. 2951–2959.
3. Deschrijver, D.; Vanhee, F.; Pissoot, D.; Dhaene, T. Automated near-field scanning algorithm for the EMC analysis of electronic devices. *IEEE Trans. Electromagn. Compat.* **2012**, *54*, 502–510.
4. Salman, H.; Ayvali, E.; Srivatsan, R.A.; Ma, Y.; Zevallos, N.; Yasin, R.; Wang, L.; Simaan, N.; Choset, H. Trajectory-optimized sensing for active search of tissue abnormalities in robotic surgery. In proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, QLD, Australia, 21–25 May 2018, pp. 1–5.
5. Berni, J.A.; Zarco-Tejada, P.J.; Suárez, L.; Fereres, E. Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 722–738.
6. Carlson, J.; Murphy, R.R. How UGVs physically fail in the field. *IEEE Trans. Robot.* **2005**, *21*, 423–437.
7. Montgomery, D.C. *Design and analysis of experiments*; John Wiley & sons: Hoboken, NJ, USA, 2017.
8. Crombecq, K.; Gorissen, D.; Deschrijver, D.; Dhaene, T. A novel hybrid sequential design strategy for global surrogate modeling of computer experiments. *SIAM JOURNAL ON SCIENTIFIC COMPUTING* **2011**, *33*, 1948–1974.
9. Van Der Herten, J.; Deschrijver, D.; Dhaene, T. Fuzzy local linear approximation-based sequential design. In proceedings of the 2014 IEEE Symposium on Computational Intelligence for Engineering Solutions (CIES), Orlando, FL, USA, 2014, pp. 17–21.
10. Galceran, E.; Carreras, M. A survey on coverage path planning for robotics. *Rob. Auton. Syst.* **2013**, *61*, 1258–1276.
11. Choset, H.; Pignon, P. Coverage path planning: The boustrophedon cellular decomposition. *Field and service robotics*. Springer: Manhattan, New York City, 1998, pp. 203–209.
12. Sadat, S.A.; Wawerla, J.; Vaughan, R. Fractal trajectories for online non-uniform aerial coverage. In proceedings of the 2015 IEEE International Conference on Robotics and Automation (ICRA), Seattle, WA, USA, 26–30 May 2015, pp. 2971–2976.
13. Johnson, M.E.; Moore, L.M.; Ylvisaker, D. Minimax and maximin distance designs. *J. Stat. Plan. Inference* **1990**, *26*, 131–148.
14. Van Steenkiste, T.; van der Herten, J.; Deschrijver, D.; Dhaene, T. ALBATROS: adaptive line-based sampling trajectories for sequential measurements. *Eng. Comput.* **2019**, *35*, 537–550. doi:10.1007/s00366-018-0614-6.
15. Jones, D.R.; Schonlau, M.; Welch, W.J. Efficient global optimization of expensive black-box functions. *J. Glob. Optim.* **1998**, *13*, 455–492.
16. Thrun, S. Exploration in active learning. *Handbook of Brain Science and Neural Networks* **1995**, pp. 381–384.
17. Auer, P. Using confidence bounds for exploitation-exploration trade-offs. *J. Mach. Learn. Res.* **2002**, *3*, 397–422.

18. Osugi, T.; Kim, D.; Scott, S. Balancing exploration and exploitation: A new algorithm for active machine learning. In proceedings of the Fifth IEEE International Conference on Data Mining (ICDM'05), Houston, TX, USA, 27–30 November 2005.
19. Raisch, S.; Birkinshaw, J.; Probst, G.; Tushman, M.L. Organizational ambidexterity: Balancing exploitation and exploration for sustained performance. *Organ. Sci.* **2009**, *20*, 685–695.
20. Williams, C.K.; Rasmussen, C.E. *Gaussian processes for machine learning*; Vol. 2, MIT press Cambridge, MA, 2006.
21. Wang, Z.; Jegelka, S. Max-value entropy search for efficient Bayesian optimization. Proceedings of the 34th International Conference on Machine Learning-Volume 70. JMLR. org, Sydney, Australia, 6–11 August 2017, pp. 3627–3635.
22. Brochu, E.; Cora, V.M.; De Freitas, N. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv preprint arXiv:1012.2599* **2010**.

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).